R & D Final Report

Contractor:	Brown University
Effective Date:	December 1, 1996
Expiration Date:	July 31, 1998
Grant Number:	F49620-97-1-0031
Principal Investigator:	Thomas Dean (401) 863-7601
Program Manager:	Capt. Alex Kilpatrick (202) 767-5028
Title:	Model Acquisition for Markov Decision Problems
Date:	October 31, 1998

Sponsored by the Air Force
Office of Scientific Research
under Contract No. F49620-97-1-0031

Final Report

The original focus of this work was on the automatic acquisition (learning) of stochastic models. The motivation was the lack of such models for military problems, specifically air-campaign planning, and the existence of new algorithms that could, if the appropriate models were available, considerably improve the accuracy and efficiency of military planning. This final report describes the course of our investigations, some unanticipated turns, and the direction that our research has taken as a consequence of what we have learned.

In recent years, we developed new models and techniques for representing stochastic processes [Dean and Kanazawa, 1988, Boutilier *et al.*, 1995a] that enabled us to compactly represent problems that couldn't be represented at all using previous techniques. We also had met with success in solving such problems using new methods that directly exploit the structure in the representations [Boutilier *et al.*, 1995b, Dean *et al.*, 1995, Dean and Lin, 1995, Lin and Dean, 1994, Lin and Dean, 1996, Lin and Dean, 1995]. Our models achieved efficiency of representation by *factoring* the state and action spaces of a dynamical system using a set of features (variously called "state variables" or "fluents"). For example, the state space for an air-campaign planning problem would have state variables for the status of each target and the location of each aircraft.

We believed when we wrote the proposal for this grant that it would be relatively straightforward to extend methods for learning hidden Markov models [Rabiner and Juang, 1986] to handle our factored representations. For certain specialized problems, researchers had already met with some success in doing exactly this [Ghahramani and Jordan, 1995]. However, in trying to carry out our research agenda [1]. we encountered two problems: First, factored models have much more structure than traditional (flat) hidden Markov models and the class of problems we were particularly interested in (highly combinatoric) was not amenable to the specialized methods in the literature. Second, in many cases, even if you could learn the models, you couldn't necessarily use the resulting representations to solve the corresponding decision problems. We found that we had some way to go in understanding the structure of factorial models and how to exploit this structure computationally before we could learn such models effectively.

Our first breakthrough came in 1997, when, in trying to understand the work of Boutilier *et al.*, we discovered how to characterize the structure their algorithm was taking advantage of in terms of bisimulation equivalence and automata equivalence [Hartmanis and Stearns, 1966]. The result was a series of papers [Dean and Givan, 1997, Givan and Dean, 1997, Dean *et al.*, 1997] in which we were able to explain the sources of combinatorial leverage

[1] We explored a wide range of approaches during the first year and carried out extensive experiments. A good deal of the material compiled during that first year is available at the Brown Computer Science Dynamics web site: http://www.cs.brown.edu/research/ai/dynamics/.

in the structured methods of Boutilier *et al.* and others. We found that the structure was due to certain symmetries in the dynamics, that, in certain cases, could be exploited to significantly reduce computation time. During the same period, we developed algorithms that were able to realize these reductions in computation time.

We also found other sources of computational leverage that were *not* accessible to these methods. In particular, we found sources of computational leverage in air-campaign planning problems that current algorithms could not handle. This prompted us to consider the sort of structure arising in systems that can be decomposed into smaller, weakly-coupled component systems. And, in 1998, we described a type of structure found in air-campaign planning problems and related logistics problems; we also developed approximation algorithms that performed extremely well on such problems [Meuleau *et al.*, 1998].

Following this unanticipated side journey, we are now returning to the problem of automatically learning stochastic models from data. We now have a great deal more experience in actually constructing (painstakingly by hand) models for air-campaign planning and related problems. We also have a much better idea of what aspects of such problems are useful to represent in the sense that they have an impact on the performance of decision-making algorithms and they provide computational leverage in solving these highly combinatoric problems. In recent months, we discovered a method for symbolically solving a system of equations of the form found in factored Markov decision processes. We also developed two structured iterative methods based on, respectively, conjugate gradient search and an acceleration method attributed to Chebyshev. These methods are of note particularly for the fact that they enable us bring to bear a large body of work on numerical methods for solving systems of equations, assuming of course that we can figure out how to factor the equations.

We are currently working on "compiler" technology that will work in concert with learning algorithms to explore the space of tractable models, rather than the much larger space of all dynamical models, many of which would do us no good even if we were to learn them. This compiler technology would enable us to identify and exploit the structure due to symmetries in the dynamics arising from (stochastic) bisimulation equivalence [Dean and Givan, 1997] and due to weakly-coupled subprocesses [Meuleau *et al.*, 1998]. We are the first to admit that this work is not traditional AI, but we are making significant progress and our approaches and methodology have been adopted by a number of labs.

References

[Boutilier *et al.*, 1995a] Boutilier, Craig; Dean, Thomas; and Hanks, Steve 1995a. Planning under uncertainty: Structural assumptions and computational leverage. In *Proceedings of the Third European Workshop on Planning.*

[Boutilier *et al.*, 1995b] Boutilier, Craig; Dearden, Richard; and Goldszmidt, Moises 1995b. Exploiting structure in policy construction. In *Proceedings IJCAI 14*. IJCAII. 1104–1111.

[Dean and Givan, 1997] Dean, Thomas and Givan, Robert 1997. Model minimization in Markov decision processes. In *Proceedings AAAI-97*. AAAI.

[Dean and Kanazawa, 1988] Dean, Thomas and Kanazawa, Keiji 1988. Probabilistic causal reasoning. In *Proceedings of the Canadian Society for Computational Studies of Intelligence*. CSCSI. 125–132.

[Dean and Lin, 1995] Dean, Thomas and Lin, Shieu-Hong 1995. Decomposition techniques for planning in stochastic domains. In *Proceedings IJCAI 14*. IJCAII. 1121–1127.

[Dean *et al.*, 1995] Dean, Thomas; Kaelbling, Leslie; Kirman, Jak; and Nicholson, Ann 1995. Planning under time constraints in stochastic domains. *Artificial Intelligence* 76(1-2):35–74.

[Dean *et al.*, 1997] Dean, Thomas; Givan, Robert; and Leach, Sonia 1997. Model reduction techniques for computing approximately optimal solutions for Markov decision processes. In Geiger, Dan and Shenoy, Prakesh Pundalik, editors 1997, *Thirteenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann.

[Ghahramani and Jordan, 1995] Ghahramani, Zoubin and Jordan, Michael 1995. Factorial hidden Markov models. In Touretzky, D. S. and Leen, T. K., editors 1995, *Advances in Neural Information Processing 7*, Cambridge, Massachusetts. MIT Press.

[Givan and Dean, 1997] Givan, Robert and Dean, Thomas 1997. Model minimization, regression, and propositional STRIPS planning. In *Proceedings IJCAI 15*. IJCAII. 1163–1168.

[Hartmanis and Stearns, 1966] Hartmanis, J. and Stearns, R. E. 1966. *Algebraic Structure Theory of Sequential Machines*. Prentice-Hall, Englewood Cliffs, N.J.

[Lin and Dean, 1994] Lin, Shieu-Hong and Dean, Thomas 1994. Exploiting locality in temporal reasoning. In Sandewall, E. and Backstrom, C., editors 1994, *Current Trends in AI Planning*, Amsterdam. IOS Press.

[Lin and Dean, 1995] Lin, Shieu-Hong and Dean, Thomas 1995. Generating optimal policies for high-level plans with conditional branches and loops. In *Proceedings of the Third European Workshop on Planning*. 205–218.

[Lin and Dean, 1996] Lin, Shieu-Hong and Dean, Thomas 1996. Exploiting locality in temporal reasoning. *Computational Intelligence* 12(3):423–449.

[Meuleau *et al.*, 1998] Meuleau, Nicolas; Boutilier, Craig; Hauskrecht, Milos; Kaelbling, Leslie; Kim, Kee-Eung; Peshkin, Leonid; and Dean, Thomas 1998. Solving very large weakly coupled Markov decision processes. In *Proceedings AAAI-98*. AAAI.

[Rabiner and Juang, 1986] Rabiner, L. R. and Juang, B. H. 1986. An introduction to hidden Markov models. *IEEE ASSP Magazine* 4–15.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	31 October 1998	1 December 1996 - 41 July 1998

4. TITLE AND SUBTITLE
Model Acquisition for Markov Decision Problems

5. FUNDING NUMBERS
F49620-97-1-0031
(G)

6. AUTHOR(S)
Thomas L. Dean

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brown University
Providence, RI 02912

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
110 Duncan Avenue Room B115
Bolling AFB DC 20332-8080

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In this research, we developed new models and techniques for representing stochastic processes that enabled us to compactly represent problems that couldn't be represented at all with previous techniques. We also developed new algorithms for efficiently solving such problems by directly exploiting the structure in the representations. Our models achieved efficiency in representation by "factoring" the state and action spaces of a dynamical system using a set of features (variously called "state variables" or "fluents"). We were able to explain the sources of combinatorial leverage in our and other structured models such as those of Boutilier et al. We found that the structure was due to certain symmetries in the dynamics that, in certain cases, could be exploited to significantly reduce computation time. Using these insights, we developed new algorithms that realized these reductions in computation time.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	4
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102